What is claimed is:

1. A method for indexing textual content in any of a plurality of languages for searching purposes, comprising the steps of:

separating a string of text into individual word tokens;

reducing the word tokens to grammatical stems by removing word endings which are associated with any one or more of the languages, without regard to whether the remaining stem is a recognized word in any combination of the plurality of languages; and

storing the stems in an index.

2. The method of claim 1 wherein the word endings which are removed are limited to only those endings which are associated with nouns.

3. The method of claim 1 wherein a word ending is not removed if the resulting stem is less than a predetermined length.

4. The method of claim 3 wherein said predetermined length is four characters.

5. The method of claim 1 wherein the reducing step is only carried out once per word token.

6. The method of claim 5 wherein said reducing step is performed by first examining each word token for the longest known endings, and examining the token for successively shorter endings until a known ending is identified in the word token and removed.

7. The method of claim 1 further including the step of disregarding stopwords during said removing and storing steps, wherein stopwords are words which occur with relatively high frequency in at least one of said languages and which are not also significant nouns in another one of said languages.

8. A method for searching for documents which may contain text in any of a plurality of languages, comprising the steps of:

separating text in each document to be searched into individual word tokens;

reducing the word tokens to grammatical stems by removing word endings which are associated with any one or more of the languages, without regard to whether the remaining stem is a recognized word in any of the plurality of languages;

storing the stems in an index which identifies the documents in which words containing the stems appeared;

receiving a query containing a string of text to be searched;

parsing the string of text into individual word tokens;

reducing the word tokens from said query to grammatical stems by removing word endings which are associated with any one or more of the languages, without regard to whether the remaining stem is a recognized word in any of the plurality of languages;

searching the index for entries which match the stems obtained from said query; and

displaying an identification of the documents which contained matching entries.

9. The method of claim 8 further including the step of displaying a matching entry along with the identification of the document in which it appears,

wherein a stem is displayed together with an ending to present a full word to the user.

10. The method of claim 8 wherein a stem is stored in said index together with the ending that was removed from a word token to form that stem, and an entry in the index that matches a stem from a query is displayed with said stored ending.

11. A system for searching for documents which may contain text in any of a plurality of languages, comprising:

a tokenizer which receives text strings from documents to be searched and user queries, and separates the text into individual word tokens;

a stemmer which reduces the word tokens to grammatical stems by removing word endings which are associated with any one or more of the plurality of languages, without regard to whether the remaining stem is a recognized word in any of the plurality of languages;

an index which stores the stems from documents and identifies the documents in which words containing the stems appeared;

a search engine which searches the index for entries which match the stems obtained from user queries; and

a display system which displays an identification of the documents which contain matching entries.

12. The system of claim 11 wherein said display system displays a matching entry from said index along with the identification of the document in which it appears, and a stem is displayed together with an ending to present a full word to the user.

13. The system of claim 12 wherein a stem is stored in said index together with an ending that was removed from a word token to form that stem, and an entry in the index that matches a stem from a query is displayed with said stored ending.

14. The system of claim 12 wherein the ending that is displayed with the stem is an ending that was removed from a word token in the query.

15. A computer-readable medium containing a program which executes the steps of:

separating a string of text into individual word tokens;

reducing the word tokens to grammatical stems by removing word endings which are associated with any one or more of the languages, without regard to whether the remaining stem is a recognized word in any of the plurality of languages; and

storing the stems in an index.

16. The computer-readable medium of claim 15 wherein the word endings which are removed are limited to only those endings which are associated with nouns.

17. The computer-readable medium of claim 15 wherein a word ending is not removed if the resulting stem is less than a predetermined length.

18. The computer-readable medium of claim 15 wherein said reducing step is performed by first examining each word token for the longest known endings, and examining the token for successively shorter endings until a known ending is identified in the word token and removed.